Protein Expression and Purification 108 (2015) 9–12

Contents lists available at ScienceDirect

Protein Expression and Purification

journal homepage: www.elsevier.com/locate/yprep

Use of the amicyanin signal sequence for efficient periplasmic expression in *E. coli* of a human antibody light chain variable domain

Brian A. Dow [a], Suren A. Tatulian [b], Victor L. Davidson [a,*]

[a] Burnett School of Biomedical Sciences, College of Medicine, University of Central Florida, Orlando, FL 32827, United States
[b] Department of Physics, University of Central Florida, Orlando, FL 32816, United States

ARTICLE INFO

Article history:
Received 24 November 2014
and in revised form 24 December 2014
Available online 5 January 2015

Keywords:
Signal sequence
Amicyanin
Protein folding
Protein expression
Antibody

ABSTRACT

Periplasmic localization of recombinant proteins offers advantages over cytoplasmic protein expression. In this study signal sequence of amicyanin, which is encoded by the *mauC* gene of *Paracoccus denitrificans*, was used to express the light chain variable domain of the human κIO8/O18 germline antibody in the periplasm of *Escherichia coli*. The expressed protein was purified in good yield (70 mg/L of culture) in one step from the periplasmic fraction by affinity chromatography using an engineered hexahistidine tag. Circular dichroism spectroscopy was used to determine if the secondary and tertiary structures of the protein and its thermal stability corresponded to those of the native folded protein. The expressed and purified protein was indeed properly folded and exhibited a reasonable thermal transition temperature of 53 °C. These results indicate that the amicyanin signal sequence may be particularly useful for prokaryotic expression of proteins which are prone to mis-folding, aggregation or formation of inclusion bodies, all of which were circumvented in this study.

Introduction

Escherichia coli is widely used in the production of recombinant proteins, including approximately 30% of therapeutic proteins that have been approved by the FDA [1]. Problems associated with protein expression in *E. coli* include degradation by intracellular proteases and the correct formation of disulfide bonds in the cytoplasm. The correct formation of disulfide bonds is especially important in biological therapeutics, where inefficient disulfide formation has been shown to limit the yield of the expressed protein [2]. High levels of cytoplasmic expression of recombinant proteins can also lead to formation of inclusion bodies and mis-folding of the proteins. This is especially prevalent during the expression of recombinant immunoglobulin domains [3–6]. These problems can be circumvented by translocation of the recombinant proteins into the periplasm of *E. coli*. The periplasm is an oxidizing environment compared to the cytoplasm. This facilitates formation of disulfide bonds that are often required for correct protein folding. The periplasm has a lower concentration of proteolytic enzymes and host cell proteins in general [7]. Thus, the periplasm is enriched in the relative concentration of the recombinant protein which facilitates its purification. For the above reasons the formation of inclusion bodies in the periplasm is much less likely than in the cytoplasm. Recombinant proteins are translocated to the periplasm via membrane-associated secretion systems that recognize an N-terminal signal sequence. The cleavage of the signal peptide during export of the recombinant protein to the periplasm yields a protein with its correct N-terminal residue. A variety of signal peptides have been used for this purpose. Some of the most widely used signal peptides are pelB from *Erwinia carotovora* and ompA, DsbA, and TolB from *E. coli* [8–10].

Amicyanin [11] is a periplasmic type I copper protein, referred to as a cupredoxin [12], which is encoded by the *mauC* gene [13] of *Paracoccus denitrificans*. Recombinant amicyanin was previously expressed in *E. coli* at high levels [14]. In this study, we describe the use of the N-terminal signal sequence of amicyanin to express the light chain variable domain (V_L)[1] of a human antibody in the periplasm of *E. coli*. The κI O8/O18 germline antibody V_L was chosen for expression as it is a well characterized member of this family [15]. Correct folding of this protein requires formation of a disulfide bond and as with many antibodies and single chain fragments, aggregation and incorrect folding of these recombinant proteins has been problematic [16]. The folding and stability of this V_L has

* Corresponding author at: Burnett School of Biomedical Sciences, College of Medicine, University of Central Florida, 6900 Lake Nona Blvd., Orlando, FL 32827, United States. Tel.: +1 407 266 7111; fax: +1 407 266 7002.
E-mail address: victor.davidson@ucf.edu (V.L. Davidson).

[1] *Abbreviations used:* V_L, light chain variable domain; CD, circular dichroism; T_m, thermal transition temperature.

http://dx.doi.org/10.1016/j.pep.2014.12.017
1046-5928/

previously been characterized by circular dichroism (CD) and thermal stability studies [15]. Thus, it was possible to assess the integrity of the recombinant protein that was expressed in this study by comparison to those results. The results of this study demonstrate the utility of the amicyanin signal sequence as an alternative for recombinant protein expression in *E. coli*. It should also be noted that an increasing number of recombinant human single chain fragments are being developed for therapeutics and so there may be additional applications of the results which are presented.

Materials and methods

Protein expression

DNA encoding the κI O8/O18 germline V_L with the N-terminal signal sequence of *mauC* from *P. denitrificans* and a C-terminal hexahistidine tag was codon-optimized for expression in *E. coli*, and synthesized by GenScript (Piscataway, NJ). The sequence of the DNA and the amino acid sequence of the protein that it encodes are shown in Table 1. The synthetic gene was cloned into a pET11a expression plasmid at the Nde1 and BamH1 restriction sites. *E. coli* BL21(DE3) cells were transformed with this plasmid, and cells were cultured in LB media at 37 °C in the presence of 100 μg/ml ampicillin. When the A_{600} of the culture reached 0.6, 0.4 mM IPTG was added to the culture to induce expression and incubation was continued at 30 °C for 12 h.

Protein purification

The recombinant κI O8/O18 V_L was purified from the periplasmic fraction of the *E. coli* cells. The periplasmic fraction was obtained using a lysozyme/osmotic shock method [11,17]. The harvested cells were resuspended in 10 mM Tris–HCl buffer, pH 8.0, at a ratio of 6 mL/g of wet cell weight. The buffer also contained 20% w/v sucrose, 0.7 mM EDTA, 2 mg/mL of lysozyme, 1 mM $MgCl_2$, 0.01 mg/mL of DNase and 200 μM phenylmethylsulfonyl fluoride. After incubation for 20 min at 30 °C with shaking, an equal volume of H_2O was added and incubation continued for a total of one hour. The spheroplasts were then removed by centrifugation and the periplasmic fraction was in the supernatant. The His-tagged κI O8/O18 V_L was purified from the periplasmic fraction by affinity chromatography using a Ni–NTA column (Qiagen). Protein concentration was determined from the absorbance of the pure protein at 280 nm using an extinction coefficient of 16,055 M^{-1} cm^{-1} that was calculated from the amino acid composition of the protein using the Expasy ProtParam tool (http://web.expasy.org/protparam/) [18].

Table 1
Sequences of the synthetic gene used to express κI O18/O8 V_L and the protein that is encoded by the gene.

ATGATTTCCGCTACCAAAATCCGCTCATGCCTCGCGGCCTGTGTCTTGGCTGCCT
TTGGAGCCACCGGAGCCCTTGCCTCGACATTCAAATGACTCAATCCCCGTCAT
CCCTGTCAGCGAGTGTCGGTGATCGCGTCACGATCACGTGCCAGGCGTCTCA
AGACATTAGCAACTACCTGAATTGGTACCAGCAGAAACCAGGTAAGGCCCC
GAAACTCTTGATCTACGACGCGTCCAATTTGGAAACAGGCGTGCCGAGTCG
CTTTAGCGGTAAGACATTCAAATGACTCAATCCCCGTCATCCCTGTCAGCGA
GTGTCGGTGATCGCGTCACGATCACGTGCCAGGCGTCTCAAGACATTAGCAA
CGACATTCAAATGACTCAATCCCCGTCATCCC**CACCATCATCATCATCAT**TAG

M I S A T K I R S C L A A C V L A A F G A T G A L A D I Q M
T Q S P S S L S A S V G D R V T I T C Q A S Q D I S N Y
L N W Y Q Q K P G K A P K L L I Y D A S N L E T G V
P S R F S G S G S G T D F T F T I S S L Q P E D I A
T Y Y C Q Q Y D N L P Y T F G Q G T K L E I K **H H H H H H**

(Top) The DNA sequence of the gene encoding κI O18/O8 V_L with the N-terminal *mauC* signal sequence and the sequence encoding the C-terminal hexahistidine tag underlined.
(Bottom) The amino acid sequence encoded from this gene is shown. The N-terminal signal sequence which is cleaved during expression and hexahistidine tag that is used for affinity purification are underlined.

Protein characterization

Size exclusion chromatography was performed with a HiPrep 16/60 Sephacryl S-300 HR column (GE Healthcare). The equilibration and elution buffer was 10 mM potassium phosphate, pH 7.5, with 150 mM NaCl. The flow rate was 1.0 mL/min. Methylamine dehydrogenase, cytochrome *c*-553, cytochrome *c*-551i, and amicyanin with molecular weights of 124, 30, 22, and 11.5 kDa, respectively, were used as standards.

CD spectra were recorded using a J-810 spectropolarimeter equipped with a Peltier temperature controller (Jasco corp., Tokyo, Japan). Samples contained 3 μM protein in 10 mM potassium phosphate buffer, pH 7.5. For temperature-dependence studies, samples were equilibrated at each temperature for 2 min before CD measurements were recorded. The measured ellipticity, θ_{meas}, was converted to mean residue molar ellipticity $[\theta]$, through the formula $[\theta] = \theta_{meas}/nlc$, where $n = 113$ is the number of amino acid residues in the protein, l is the optical pathlength in mm, and c is the molar concentration of the protein. Thermal unfolding of the protein secondary structure was determined from the increase in the negative CD signal at 205 nm which results from an increase in the fraction of unordered structure [19]. The value of the thermal transition temperature (T_m) was determined as the inflection point of the sigmoidal temperature-dependence curves of the corresponding normalized signal intensity, which was identified as the extremum of the first derivative with respect to the temperature. The fraction of folded protein at a given temperature T was determined using Eq. (1) where $[\theta]$ is the mean residue molar ellipticity at 205 nm and the subscript indicates the temperature. Values of $[\theta]$ at 40 °C and 70 °C were deemed to correspond to 100% and 0% folded state, respectively.

$$\text{Fraction folded} = \frac{[\theta]_T - [\theta]_{70}}{[\theta]_{40} - [\theta]_{70}} \quad (1)$$

Results

Protein expression and purification

The recombinant κI O18/O8 germline V_L was successfully expressed in *E. coli* and purified. The yield of pure protein was 70 mg/L of cell culture. The purified protein ran as a single band on SDS–PAGE (Fig. 1A). When the protein was subjected to size exclusion chromatography it eluted as a single peak consistent with it being present exclusively as a monomer (Fig. 1B). Comparison of the elution volume of the V_L to other protein standards indicated that it eluted at a position corresponding to a molecular mass of about 9 kDa which is lower than that determined by SDS–PAGE and calculated from the amino acid sequence of 12.8 kDa. However, this observation is consistent with anomalously lower values for molecular mass determined by size exclusion chromatography that have been previously observed with V_L domains [20–22]. Since the expressed protein is an antibody domain and not an enzyme, it does not have an activity to assay. Evidence for the structural and functional integrity of the expressed protein was provided by the demonstration that it binds to protein L resin, which is specific for antibody kappa light chains, and is often used for affinity purification of immunoglobulins [23]. In this study we chose to add a His-tag for purification because the Ni–NTA superflow resin used for purification has a much higher capacity for bound protein than the protein L resin. It was also confirmed by Western blot that the purified protein reacted with an anti-His-tag antibody. These results show that use of the amicyanin signal sequence allowed